

Automated Translation Software Opening Markets

November 17, 2005

By Dave Pelland, Managing Editor, Technology Insider

As globalization increases the need to overcome language barriers, multinational companies are enlisting automated translation software to help evaluate new markets, provide customer support and monitor international news events.

Computerized translation -- known as automated or machine translation -- has been an industry goal since the earliest days of computer science. American military and intelligence agencies have been exploring the technology for nearly 30 years.

While not as refined as human translations, statistics-based methods developed in the past five years produce translations that allow readers to get the basic meaning of documents or speech produced in another language.

"With statistical products coming onto the market, you're seeing a higher quality of translation," says Bryce Benjamin, chief executive of software provider Language Weaver.

Companies in numerous industries are using translation software to prepare sales materials and support documents, as well as to evaluate conditions in emerging markets.

Until recently, technology attempted to match words in a source language with definitions in the target language. But because words can have different meanings in different contexts, it's difficult to create rules that produce accurate translations.

"A lot of times with machine translation, you don't know if the translation is good or bad because you don't understand the other language," says Dina Moscovitz, vice president of marketing and communications for Spoken Translation Inc. "And even if you know it wasn't a good translation, you wouldn't know how to fix it on your own."

For example, the word "bank" can refer to one side of a river, an airplane turn or a financial institution. Statistics-based translation software examines other words with a sentence or paragraph, such as "deposit" or "canoe," to help determine the proper context.

Language Weaver software starts with millions of documents that have previously been translated between two languages, as well as Web sites, news reports and other data. The software "learns" statistical correlations between words and phrases in both languages, and how words are strung together.

"We're looking at all the data and calculating the probability that a word string in the source language matches a word string in the target language," Benjamin says.

A resulting translation is unlikely to match the quality of a human translation, but is less expensive and faster to prepare.

Translation software is typically sold on an enterprise basis. Pricing varies according to the needed language pairs, in part because more obscure languages may require humans to prepare documents needed to help train the software. List prices for ongoing licenses and support can range between \$20,000 and \$150,000.

"[Machine] translations are not refined or edited, but if people are trying to monitor a certain subject, it's close enough," says Dimitris Sabatakakis, chairman and CEO of software provider Systran S.A. "Most of the use of our software is getting the gist while accessing foreign-language content."

The software can also help technology companies reduce support costs by allowing customers who don't speak English to service their own accounts. For example, a hardware vendor reports 80 percent of its customers in English-speaking countries access their accounts on the Web. That rate falls to under 3 percent in Asian countries, in large part because the company's online documentation is primarily in English.

The technology is also used heavily within the automotive industry to translate sales materials as well as to customize owners' manuals for international markets.

A similar application, currently in its early stages, automatically translates text-based chat sessions between customers and technical support personnel. Even if the translations are imperfect, they can help customers with an underlying problem.

Automated translation is also being used in health care settings. For instance, Spoken Translation is introducing software for tablet PCs designed to promote English-Spanish communication. A doctor or nurse can ask questions that are displayed on the screen, and the patient can respond using a keyboard or handwriting recognition.

Moscowitz says the software isn't intended to replace a human translator in an emergency department, or meet all of health care providers' translation compliance requirements, but to provide an affordable alternative for routine settings.

Language Weaver's Benjamin says the software is also being used by companies and homeland security agencies to monitor international news reports. In one application, Chinese-language broadcasts are being treated with speech-recognition software to prepare transcripts that, in turn, can be translated into English.

"Some multinational companies are monitoring events on a daily basis that can affect their business or employees," Benjamin says. "That mirrors what we see in the government, which follows a lot of media around the globe."

In many cases, existing translations can help expand industry-specific data used in comparing documents.

"Companies that have been doing a lot of multinational business for a long time already have lots of translated information in the languages that are of interest to them with all the industry terminology and jargon," Benjamin says. "We can use that as training material."

Sabatakakis says another factor promoting automated translation is the electronic preparation and distribution of documents, which reduces the need to scan printed materials. Gains in raw computing power and reductions in the cost of storing large amounts of translated data are also helping the technology.

In addition, the use of standardized formats and XML to tag data fields within electronic documents increase the efficiency of automated translation, because the software doesn't have to search for trademarked words that don't have an equivalent in another language.

"When a company can specify items like product names [in electronic documents], there's no ambiguity," Sabatakakis says.

As translation software improves, one of the challenges vendors face is tempering customers' expectations about the ease and quality of automated translations.

"Everybody dreams of clicking a button and having millions of documents translated, so it's very important to explain what's possible," Sabatakakis says.